

Boosting the Area Under the ROC Curve (Poster M32)

Phil Long (Google) and Rocco Servedio (Columbia)

- Ranking problem
 - Given examples of members and nonmembers of a class, output scoring function
 - Goal: highest scorers mostly members
- AUC (Area Under the ROC Curve)
 - Standard criterion
 - Value $\frac{1}{2}$ for random output
 - Value 1 for perfect output
- Theorem
 - Can “boost” the AUC (slightly better than $\frac{1}{2}$ implies close to 1)
 - ... also with misclassification noise at rate $\eta < \frac{1}{2}$.
 - Contrasts w/ classification, where $1 - \eta$ best possible [KS05]

